

Some Nonlinear Phenomena in the Ionosphere

V. A. Bailey¹

Emeritus Professor, University of Sydney, Sydney, Australia

(Received May 4, 1964)

The historical development is described of the idea that sufficiently strong radio waves can modify the properties of the ionospheric regions which they traverse.

The modified primary properties are the electron temperature and collision frequency, the rates of electron attachment and recombination, and the excitation and ionization of molecules.

The resulting changes in the collision frequency and electron density modify the region's absorption coefficient and refractive index for any wave, including the modifying wave itself. These lead to the phenomena of wave interaction and self-distortion, respectively.

The associated mathematical theory includes the important effects on gyrowaves due to the presence of the earth's magnetic field. The principal predictions of this theory and their subsequent experimental verification are discussed.

Applications of this theory to the study of the ionosphere and laboratory plasmas are indicated. These include control of the ionosphere and the generation of artificial airglows (or aurorae) by means of powerful gyrowaves.

The possible modification of the ionosphere by low-frequency extraterrestrial radio waves is also considered; such waves include geomagnetic disturbances. They may arise from hydromagnetic disturbances in the solar wind and from fluctuating screening by plasma clouds of postulated electric charges on the sun.

1. Introduction

In a letter to *Nature*, B. D. H. Tellegen [1933] reported that the transmissions of the Swiss radio station at Beromunster (650 kc/s) received in Holland appeared to be modulated in the ionosphere by the radiation from the newly erected powerful station at Luxembourg (200 kW, 252.1 kc/s). This and an earlier report by Butt [1933] are examples of the phenomenon of ionospheric wave interaction² which was subsequently a source of annoyance to European broadcast listeners. The observations reported by B. van der Pol [1935] and others all indicated that a powerful station of sufficiently low frequency can impress its modulation on the wave received from another station (the wanted wave) when the latter traverses ionospheric regions whose projections on the ground lie within about 100 to 200 km of the disturbing station. The two waves are now commonly called the *disturbing wave* and the *wanted wave*. The observed depth of impressed modulation could amount to about 10 percent.

It is worthy of note by applied physicists that the accepted theory of this phenomenon arose exclusively from a "thought experiment" which occurred to the present author while officially examining a doctoral candidate's exposition of the magneto-ionic theory of ionospheric waves. This theory was first mentioned in a letter to *Nature* [Bailey and Martyn, 1934] and more precisely

developed in subsequent publications [Bailey, 1937a, b, c, d; 1938a, b]. A brief account of this development is given in a Supplement to *Nuovo Cimento* [Bailey, 1956].

Since most of the experimental work on wave interaction has been guided by these mathematical developments, it is here desirable to present the theory first.

The two fundamental concepts on which the theory is based are as follows:

1. Any electric field of intensity E (V/cm) applied to free electrons in thermal equilibrium with a gas at pressure p (mm Hg) and temperature T_0 ($^{\circ}$ K) will impart energy to the electrons and so raise their temperature to a higher value T . This results in an increase in the rms value u of the random velocities of the electrons and so to a change in ν the mean frequency of collision of an electron with molecules of the gas; in general this change is an *increase* of the collision frequency.

2. The absorption coefficient κ of a radio wave which traverses a particular region of the ionosphere varies with the value of the local collision frequency ν . Except for an extraordinary wave of frequency near the local gyrofrequency, κ increases³ with ν .

It follows from 1 that a powerful radio wave will in general increase the collision frequency in some region of the ionosphere, and then from 2 that as a result another wave (not an extraordinary gyro-wave) traversing this region becomes more absorbed by the region.⁴ This increased absorption of the

¹ Deceased. Send communications and reprint requests to J. Bailey, 4 Ch. de la Golette, Meyrin, Geneva, Switzerland.

² Also known as the "Luxembourg effect" and the "Tellegen effect."

³ For the exceptional extraordinary gyrowave, κ decreases as ν increases.

⁴ An extraordinary gyrowave will become less absorbed by the region.

wanted wave will reveal itself (in general) as a decrease in the amplitude of the wave at reception. Thus any modulation of the amplitude of the disturbing wave at a frequency n and to a depth M will impress a modulation of the same fundamental frequency on the wanted wave.

The same concepts 1 and 2 lead to the conclusion that a sufficiently powerful wave will modify the absorption of the regions which it traverses and so be received in a distorted form. This phenomenon has been termed "self-distortion."

Besides explaining the observations of Tellegen, Butt, and others, the theory has predicted the following phenomena:

(1) The amplitude of the wanted wave is modulated by a fraction f_i which is given by formulas (44) and (45) and consists of a fundamental part of frequency n , phase lag φ_1 and modulation depth M_1 proportional to M and a harmonic part of frequency $2n$, phase lag φ_2 , and depth M_2 proportional to M^2 .

(2) The depths M_1 and M_2 of the two impressed modulations decrease in specified manners as the modulation frequency n of the disturbing wave increases.

(3) The depths M_1 and M_2 increase nearly proportionally to the mean power \bar{P} of the disturbing station.

(4) The phase lags φ_1 and φ_2 are equal to the inverse tangents of the ratios $n/G\nu_0$ and $2n/G\nu_0$, respectively, where ν_0 is the value of the electron collision frequency in the unperturbed local ionospheric region and G is a particular constant derived from laboratory data for air; $G \sim 2 \times 10^{-3}$.

(5) The depths M_1 and M_2 decrease as the angular radio frequency ω of the disturbing wave increases within certain limits.

(6) Each of the depths M_1 and M_2 passes through one or two maximum values as the angular radio frequency ω of the disturbing wave is varied about the local angular gyrofrequency Ω , i.e., resonance occurs. Also a radiated power of 1 or 2 kW would then suffice to produce observable interaction with an appropriate wanted wave.

(7) The depths M_1 and M_2 increase as the region of ionospheric reflection of the wanted wave approaches to about 100 km from the ionospheric region vertically above the disturbing station.

(8) A pulsed disturbing wave causes the amplitude of the wanted wave to rise and decay approximately according to a particular law which involves the time constant $t_0 = 1/G\nu_0$.

(9) The observed modulation of a wave becomes increasingly distorted, at the expense of the lower modulation frequencies, as its power increases, i.e., self-distortion increases with the power.

(10) "Interaction atmospherics" can be impressed on a wanted wave by the low-frequency radiation from suitably located electric storms.

(11) Combination tones can be impressed on a wanted wave by two disturbing waves with radio frequencies which differ by an audio frequency.

The predictions (1) to (9) have been confirmed by subsequent experimental observations. Also (4) has been used by Ratcliffe and Shaw [1948] to determine the height at which a particular wave interaction occurs and (8) has been used to determine the quantity $G\nu$ which relates in part to the mean energy lost by an electron at a collision with an ionospheric molecule. The prediction (11) appears to account for certain observations of Gill [1935].

Also it may be mentioned that the formula for M_1 yielded by the theory [and given below under (45)] correctly accounts for the order of magnitude of the depth of the modulation impressed on Beromunster's wave by the Luxembourg station as observed by van der Pol and van der Mark [1935].

The fundamental concepts 1 and 2 have also led to the following predictions:

(12) A specified vertical beam of extraordinary gyrowaves with radiated power of 500 kW can generate a visible glow discharge or artificial "aurora" in the E -region at night [Bailey, 1938a, b; 1961].

(13) A vertical beam of extraordinary gyrowaves radiated with a power of 500 kW from an array of 40 horizontal dipoles can penetrate into the nocturnal E -region much more easily than a weak beam. Also, this beam suffices to increase the electron density near the 92 km level of the nocturnal E -region by an observable amount and thus to determine the frequency of attachment of electrons to molecules near this level [Bailey, 1959].

These results and the corresponding experimental studies by Goldstein et al. [1953a, b], and Anderson and Goldstein [1955], of the effects of strong microwaves on gaseous plasma have led to a proposal for control of the ionosphere by means of radio waves [Bailey and Goldstein, 1958].

We shall now consider in detail according to the following plan the principal theoretical and experimental results of investigations on ionospheric wave interaction published up to February 1963.

Section 2 summarizes the relevant concepts, methods, and experimental observations of the Townsend School, which relate to the collisions of electrons with gas molecules or atoms, under the influence of electric fields of effective intensity E V/cm at gas pressures p mm Hg and gas temperatures T_0 (usually about 15 °C).

Section 3 summarizes the relevant concepts and experimental observations, of the Appleton School and other ionospheric physicists, which relate to the propagation of radio waves in the ionosphere regarded as a magneto-ionic gaseous medium.

Section 4 briefly sets out the mathematical theory of wave interaction as a logical consequence of the work described in sections 1 to 3.

Section 5 considers in detail the principal predictions to which this theory leads.

Section 6 describes the experimental tests and verifications of these predictions.

Sections 7 and 8 discuss the use of the theory as an instrument for the study of the ionosphere and electrical gas discharges, some of its practical applications, and some recent developments of the subject.

2. Collisions of Electrons With Gas Molecules Under the Influence of an Electric Field

In the absence of an electric field, the mean energy U of an electron in a gas at uniform temperature T_0 is equal to the mean energy U_0 of a molecule⁵ of the gas.

Townsend and his collaborators [Townsend, 1947; Healey and Reed, 1940] discovered that in the presence of an electric field E , whether constant or alternating [Townsend, 1932], the mean energy of random motion U of an electron is greater than U_0 , and that for a given gas, the ratio

$$k = U/U_0$$

is a function of the ratio E/p alone, where p is the gas pressure.

By measuring, in a constant field E , the divergence in the gas of a stream of electrons which passes through a slit, they were able to determine the values of the Townsend factor k corresponding to different values of E/p . Then by deflecting the same stream through a known angle by means of a transverse magnetic field, they were able to measure the corresponding values of the drift velocity W of the electrons in the direction of the electric field.

From these known values of k and W and by means of simple⁶ approximate theoretical formulae, they then calculated the mean free path L of an electron of energy U in the gas at the standard pressure of 1 mm Hg and the mean proportion λ of its own energy which such an electron loses at a collision with a molecule.

From the value of L corresponding to a given value of U , the mass m of an electron, and the relation $U = \frac{1}{2}mu^2$, we can then calculate in succession the corresponding mean random velocity u and collision frequency ν at a pressure p by means of the formulae

$$u = (2U/m)^{1/2} = (2kU_0/m)^{1/2} \quad (1)$$

$$\nu = u/l = up/L, \quad (2)$$

since the mean free path l at the pressure p is given by

$$l = L/p.$$

Also, from the corresponding value of λ we can calculate η , the mean energy lost by an electron at a collision, by means of the formula

$$\eta = \lambda U. \quad (3)^7$$

The values of k , W , L , and λ corresponding to different values of E/p for well-dried air were first determined by Townsend and Tizard [1913]. These are also set out in the books by Townsend [1947] and Healey and Reed [1940]. More recent determinations of some of these quantities have been made by Bailey [1925], Neilsen and Bradbury [1937], and Crompton, Huxley, and Sutton [1953]. On account of the limitations set by the experimental methods used, the lowest value of the Townsend factor measured in air was $k=4$; this corresponds to a mean electron energy of 4/27 electron volts. Since the variations of L and η with U for $k < 4$ are not known, the application of the known experimental data to the lower ionosphere makes it necessary to extrapolate these data in some simple way to the range in which $1 \leq k < 4$. For this purpose the simplest two hypotheses to adopt are:

(1) L is independent of U and

(2) $\eta = G(U - U_0), \quad (4)$

where G is a constant determined by the data obtained with the lowest values of the field E . For air, the value $G \sim 2 \times 10^{-3}$ has proved the most successful.

From the first hypothesis and (1) and (2), it follows that the collision frequency ν in air always *increases* with the intensity of the field E in the range of energies considered.

Following Townsend's work on the uniform column in high frequency electrical discharges through gases ("Electrical Discharges," Llewellyn Jones [1953]) the formulae (1) to (4) may be applied to the effects caused by a strong radio wave in the ionosphere. But the dependence of these effects on the field intensity \mathbf{E} , the radio frequency $\omega/2\pi$, and the terrestrial magnetic field \mathbf{H} needs further consideration, such as that given below in section 4 and the appendix. This involves the determination of \bar{w} , the mean work done on an electron over a free path or, alternatively, the mean power $\Xi = N\nu\bar{w}$ expended on the N electrons in unit volume.

3. Wave Propagation in a Gaseous Magneto-Ionic Medium

In Appleton's well-known magneto-ionic theory of wave propagation, the medium is a gas which contains N electrons/cm³ and is pervaded by a uniform, constant magnetic field \mathbf{H} ; also the mean collision frequency ν of an electron in the gas is taken as independent of the amplitude of the wave.

The complex refractive index $M = \mu - i\chi$ and the polarization number R , which define a mode of propagation of a wave of angular frequency ω in this medium, are determined by means of Maxwell's field equations combined with the equation of transport of mean momentum, which governs the momentum $Nm\mathbf{v}$ of the electrons in unit volume. When

⁵ In this paper the term "molecule" is intended to include "atom."

⁶ The more complicated formulae proposed later by various workers have not been found necessary in the present state of experimental knowledge of wave interaction.

⁷ η can also be obtained directly from the drift velocity W by means of the formula $\eta = 1.23 \text{ mW}^2$.

N is uniform and constant, the latter reduces to the equation

$$m \frac{d\mathbf{v}}{dt} + \nu m \mathbf{v} = e \mathbf{E} + \frac{e}{c} \mathbf{v} \times \mathbf{H}, \quad (5)$$

where e is the charge on an electron in e.s. units and the magnetic vector of the wave field is neglected by comparison with \mathbf{H} .

From the dispersion equation yielded by (5) and the field equations, it follows that there are two wave modes and that for each the corresponding ray becomes attenuated as it progresses, through a factor a_s defined by

$$a_s = \exp \left(- \int_0^s \kappa ds \right) \quad (6)$$

where s is the length of the ray path measured from its beginning and κ is related to χ as follows:

$$\kappa = \omega \chi / c. \quad (7)$$

The value of the coefficient of absorption κ depends on ν , N , and the angular gyrofrequency Ω .

The theory shows that when

$$\omega^2 \gg \nu^2 \quad (8)$$

κ is in general approximately proportional to the product νN . For the exceptional case, in which the angular wave frequency ω is nearly equal to Ω and the mode is extraordinary, κ is given by (64) and so is nearly proportional to $N\nu / \{ \nu^2 + (\omega - \Omega)^2 \}$. It follows that with increasing ν the attenuation in general increases, but decreases for an extraordinary wave mode with a frequency near the gyrofrequency. The latter may be conveniently termed an extraordinary gyrowave.

In a slowly varying nonuniform medium a ray theory of propagation may be used. Then along a ray path of length s , the field amplitude is given by

$$E = E_0 \exp(-\alpha)$$

where

$$\alpha = \int_0^s \kappa ds.$$

Since the mean power flux density is \bar{P} where

$$\bar{P} = c \mu \bar{E}_r^2 / 4\pi,$$

where \bar{E}_r is the rms value of the component of \mathbf{E} transverse to the direction of propagation, it follows that

$$\bar{P}/\mu = (\bar{P}_0/\mu_0) \exp(-2\alpha) \quad (9)$$

and so

$$\partial(\bar{P}/\mu) / \partial s = -2\kappa(\bar{P}/\mu). \quad (10)$$

4. Mathematical Theory of Ionospheric Wave Interaction

From consideration of the results mentioned in section 2, we see that in (5) the quantity ν is not really a constant, but must now be regarded as a function of the wave field intensity.

It follows that the classical magneto-ionic formulae for M and R must now be regarded as first approximations which are accurate only when the waves concerned are weak enough to make ν closely equal to the normal collision frequency ν_0 which exists in the unperturbed medium.

The fact that at least κ is a function of the field intensity E is sometimes referred to as the "non-linearity" of the medium.⁸

In order to obtain higher approximations to M and R it is in general necessary to apply lengthy perturbation procedures. So in order to keep the present mathematical theory sufficiently simple for practical purposes we shall relax its rigor to the limits allowed by the uncertainties in the physical data involved.

4.1. Work Done on an Electron by a Plane-Polarized Wave Over a Mean Free Path

It has been shown [Bailey, 1937b] that

$$w = \frac{e^2}{2m} E_e^2 [A + A \cos 2\omega t - B \sin 2\omega t] \quad (11)$$

where E_e is the effective value $E/\sqrt{2}$ of the local electric field intensity \mathbf{E} ,

$$\left. \begin{aligned} A &= \frac{s^2}{\nu^2 + (\omega - \Omega)^2} + \frac{s^2}{\nu^2 + (\omega + \Omega)^2} + \frac{2c^2}{\nu^2 + \omega^2} \\ B &= - \left(\frac{\omega - \Omega}{\nu^2 + (\omega - \Omega)^2} + \frac{\omega + \Omega}{\nu^2 + (\omega + \Omega)^2} \right) \frac{s^2}{\nu} - \frac{2\omega}{\nu^2 + \omega^2} \frac{c^2}{\nu} \end{aligned} \right\} \quad (12)$$

ω is the angular frequency of the wave,

$\Omega = He/mc$ (angular gyrofrequency),

$s = \sin \psi$, $c = \cos \psi$

and ψ is the angle between \mathbf{E} and \mathbf{H} .

The mean work \bar{w} done during a complete cycle $2\pi/\omega$ and over a free path is therefore given by

$$\bar{w} = \frac{e^2}{2m} E_e^2 A. \quad (13)$$

In the following sections we shall assume that E_e varies appreciably only during a period of time which is large compared with the period $2\pi/\omega$ of the wave.

⁸ The experiments of King, described in section 6, show that notable ionospheric self-distortion can occur when the radiated power of the transmitter is at least 200 kW.

4.2. Differential Equation for the Mean Kinetic Energy U of an Electron

The equation of transport of energy in the electron gas is

$$\frac{d(NU)}{dt} + N\nu\eta = N\nu w,$$

so when the wave is not strong enough to change N appreciably we have

$$\frac{dU}{dt} + \nu\eta = \nu w. \quad (14)$$

When the energy U does not exceed $4U_0$ we have $1 < k \leq 4$ and so we may adopt the formula (4) as a working hypothesis. Then (14) leads to the equation

$$\frac{dU}{d\tau} + (\nu/\nu_0)(U - U_0) = t_0\nu w \quad (15)$$

where

$$\left. \begin{aligned} t_0 &= 1/G\nu_0 \\ \tau &= t/t_0 \end{aligned} \right\}. \quad (16)$$

t_0 is a time constant and τ is the time measured in units equal to this time constant.

Also, by (2) and (1)

$$\nu/\nu_0 = (U/U_0)^{1/2} (L_0/L). \quad (17)$$

The dependence of L on U for $U < 4U_0$ is not yet known with certainty. But the experiments on air of Townsend and Tizard and of Crompton, Huxley, and Sutton indicate that no great error is likely to result from our adopting here the original hypothesis that L is independent of U , i.e.,

$$\nu/\nu_0 = (U/U_0)^{1/2}. \quad (17.1)$$

By substitution for ν from (17) or (17.1) in (12) and from (11) we can express νw as a known function of U and E_e^2 . Then (15) becomes a differential equation in the unknown function U with E_e as a given slowly varying function of t , of the form

$$E_e = E_{e1}M(t) = E_{e1}M(t_0\tau) \quad (18)$$

where E_{e1} is a constant field. $M(t)$ may be termed the modulation function or modulation envelope.

The resultant differential equation requires some method of approximation for its solution. From such a solution U and (17) or (17.1) we can then obtain an approximate expression for ν as a function of t , E_{e1} , ω , Ω , ψ , G , and U_0 or ν_0 .

These expressions for U and ν contain terms which vary slowly with t and terms which are periodic in t with 2ω as their fundamental frequency. Retention of the latter terms in the discussion of wave interaction and self-distortion would lead to similar periodic terms in these phenomena. But at radio frequencies ω such terms are usually negligible by comparison with the slowly varying terms, so henceforth they will be neglected.

Accordingly we will now consider only the variations of U and ν which take place in a time dt large compared with the wave period $2\pi/\omega$. This requires the symbol w in (14) and (15) to be replaced always by its mean value \bar{w} which is given by (13) or by (48).

4.3. Differential Equation for the Mean Collision Frequency ν

It is now convenient to base the discussion on the equation for ν . Thus, on eliminating U from (15) by means of (17.1), replacing w by \bar{w} , and using (13) and (18), we obtain

$$\frac{d\sigma}{d\tau} + \frac{1}{2}(\sigma^2 - 1) = \bar{w}/2GU_0 = RM^2(t_0\tau) \quad (19)$$

where

$$\sigma = \nu/\nu_0, \quad (20)$$

$$R = \frac{e^2}{4GmU_0} E_{e1}^2 A. \quad (21)$$

Here σ is the collision frequency expressed in units equal to ν_0 .

On account of its factor A , the quantity R in (19) is a function of σ^2 such that this equation has no obvious solution in finite terms.

However, in order to apply the theory to most of the observations of wave interaction which have been made hitherto, it is sufficient to consider only those approximate solutions which correspond to values of $\bar{w}/2GU_0$ or $RM^2(t_0\tau)$ such that σ varies about some constant value σ_1 by a small fraction of σ_1 , i.e.,

$$\left. \begin{aligned} \sigma &= \sigma_1 + \delta, \\ \sigma_1 &= \nu_1/\nu_0, \quad |\delta| \ll \sigma_1 \end{aligned} \right\}. \quad (22.1)$$

where

Thus

$$\nu = \nu_1 + \nu_0\delta, \quad |\nu_0\delta| \ll \nu_1. \quad (22.2)$$

These approximate solutions are then found from solutions of the equation:

$$\frac{d\delta}{d\tau} + \sigma_1\delta = R_1M^2(t_0\tau) + \frac{1}{2}(1 - \sigma_1^2) \quad (23)$$

where

$$R_1 = \frac{e^2}{4GmU_0} E_{e1}^2 A_1 \quad (21.1)$$

and

$$A_1 = \frac{s^2}{\nu_1^2 + (\omega - \Omega)^2} + \frac{s^2}{\nu_1^2 + (\omega + \Omega)^2} + \frac{2c^2}{\nu_1^2 + \omega^2}. \quad (24)$$

4.4. Collision Frequency ν_s in the Steady State Under the Influence of a Continuous Wave

Here we have $M^2(t_0\tau) = 1$. Then the steady value $\sigma_1 = \nu_s/\nu_0$ is obtained by setting $d\sigma/dt = 0$ in (19) and solving the resultant quartic in σ^2 or by setting $\delta = 0$ in (23) and solving the same quartic in σ_1^2 .

When E_{e1} is not too large,⁹ we then deduce the following approximate expression for the steady collision frequency:

$$\nu_s = \nu_0 + R_0 \nu_0, \quad (25)$$

where

$$R_0 = \frac{e^2}{4GmU_0} E_{e1}^2 A_0 \quad (26)$$

and A_0 is the same as A_1 in (24) with ν_1 replaced by ν_0 .

From (25), (26), and (24) it follows that, except when the wave's field vector \mathbf{E} is only slightly inclined to the constant magnetic field vector \mathbf{H} , ν_s passes through a maximum value as the wave frequency ω is varied around the gyrofrequency Ω . The corresponding mean electron energy U_s behaves similarly.

These resonance effects and their consequences are of great importance, as will be shown below.

4.5. Rise and Decay of ν Caused by a Square Pulse of Radio Waves

When $\nu = \nu_0$ at $t=0$ and $M^2(t_0\tau)=1$ for $0 \leq t \leq T$ then (23) has $\sigma_1=1$ and its right-hand side constant.

The corresponding solution of (23) which describes the rise of ν , is

$$\nu = \nu_s - (\nu_s - \nu_0) e^{-t/t_0}. \quad (27)$$

Similarly we obtain the following solution which describes the decay of ν from the steady value ν_s :

$$\nu = \nu_0 + (\nu_s - \nu_0) e^{-t/t_s} \quad (28)$$

where

$$t_s = 1/G\nu_s.$$

For the E -region t_0 is of the order of 1 msec or less. Hence the rise and decay require about 1 msec to be practically completed.

Accurate expressions for the rise and decay can also be obtained, since (19) is now exactly integrable [Bailey, 1937b, eq (34)].

4.6. Modulation of ν Caused by a Wave Which Is Sinoidally Modulated in Amplitude

For such a wave we have

$$\left. \begin{aligned} M(t) &= 1 + M \sin nt \\ 0 < M < 1, \quad n < \omega \end{aligned} \right\} \quad (29)$$

Then

$$\begin{aligned} M^2(t_0\tau) &= 1 + \frac{1}{2}M^2 + 2M \sin(nt_0\tau) \\ &\quad - \frac{1}{2}M^2 \cos(2nt_0\tau). \end{aligned} \quad (29.1)$$

If we take

$$\nu_1 = \nu_s$$

for the steady state corresponding to the mean effective force

$$E_{e1}(1 + \frac{1}{2}M^2)^{\frac{1}{2}}$$

then the constant part of the R.H.S. of (23) vanishes, i.e.,

$$R_1(1 + \frac{1}{2}M^2) + \frac{1}{2}(1 - \sigma_1^2) = 0, \quad (29.2)$$

and so (23) reduces to

$$\frac{d\delta}{d\tau} + \sigma_1\delta = 2R_1M \sin(nt_0\tau) - \frac{1}{2}R_1M^2 \cos(2nt_0\tau), \quad (30)$$

where

$$\sigma_1 = \nu_s/\nu_0,$$

and, by (25),

$$\nu_s = \nu_0 + \frac{e^2}{4GmU_0} E_{e1}^2 (1 + \frac{1}{2}M^2) A_0 \nu_0. \quad (31)$$

The general solution of (30) combined with (22.2) leads to the following formula for ν :

$$\nu = \nu_s + \frac{BM \sin(nt - \varphi_1)}{(1 + n^2 t_s^2)^{\frac{1}{2}}} - \frac{\frac{1}{4}BM^2 \cos(2nt - \varphi_2)}{(1 + 4n^2 t_s^2)^{\frac{1}{2}}} + Ce^{-t/t_s} \quad (32)$$

where

$$\left. \begin{aligned} B &= \frac{e^2 E_{e1}^2 A_s t_s}{m^2 t^2} \\ t_s &= 1/G\nu_s \\ \tan \varphi_1 &= nt_s, \quad \tan \varphi_2 = 2nt_s, \end{aligned} \right\} \quad (33)$$

A_s is the same as A_1 in (24) with ν_1 replaced by ν_s and C is a constant determined by the condition $\nu = \nu_0$ at $t_0 = 0$.

4.7. Modulation of a Wanted Wave Through the Variation of the Local Collision Frequency

In passing through the ionosphere, the field amplitude F of the wanted wave mode is attenuated according to the formula

$$\left. \begin{aligned} F &= F_0 \exp(-\alpha) \\ \alpha &= \int_0^s \kappa(\nu) ds, \end{aligned} \right\} \quad (34)$$

where

s is the length of the wanted wave's path measured from its entry into the ionosphere, F_0 is the field amplitude at entry, and $\kappa(\nu)$ is the absorption coefficient corresponding to the value ν of the collision frequency at the point s in the path.

Under those conditions in which the collision frequency varies about some constant value ν_1 , i.e.,

$$\nu = \nu_1 + \Delta\nu \quad (35)$$

we have

$$\alpha = \alpha_1 + \Delta\alpha \quad (36)$$

⁹ More definitely when $R_0 \ll \frac{1}{2}$.

where

$$\alpha_1 = \int_0^s \kappa(\nu_1) ds$$

and

$$\Delta\alpha = \int_0^s \{ \kappa(\nu) - \kappa(\nu_1) \} ds. \quad (37)$$

Hence, by (34) and (35)

$$F = F_1 \exp(-\Delta\alpha)$$

where

$$F_1 = F_0 \exp(-\alpha_1).$$

So at the point s the amplitude of the wanted wave is modulated according to the modulation function $M_i(t)$ where

$$M_i(t) = \exp(-\Delta\alpha).$$

If we set

$$M_i(t) = 1 + f_i(t) \quad (38)$$

the impressed modulation fraction $f_i(t)$ is given by

$$f_i(t) = \exp(-\Delta\alpha) - 1.$$

When observation shows that $|f_i(t)| \ll 1$, we will necessarily have the approximation

$$f_i(t) = -\Delta\alpha = - \int_0^s \{ \kappa(\nu) - \kappa(\nu_1) \} ds. \quad (39)$$

On substituting for ν from (35) and expanding $\kappa(\nu_1 + \Delta\nu)$, we obtain

$$f_i(t) = - \int_0^s \kappa'(\nu_1) \Delta\nu ds \quad (40)$$

provided that

$$|\Delta\nu \kappa''(\nu_1) / \kappa'(\nu_1)| \ll 2.$$

When $RM^2(t)$ is sufficiently small for (22.2) to hold true, we have, by (35),

$$\Delta\nu = \nu_0 \delta$$

where δ satisfies the equation (23). So, on multiplying (23) throughout by $-\kappa'(\nu_1)\nu_0 ds$, integrating with respect to s , and using (40) we obtain the following equation for $f \equiv f_i(t)$, for a situation in which σ_1 is nearly constant throughout the region concerned:

$$\frac{df}{d\tau} + \sigma_1 f = S_1 M^2(t_0\tau) + S_2 \quad (41)$$

where

$$\left. \begin{aligned} S_1 &= - \int_0^s R_1 \kappa'(\nu_1) \nu_0 ds \\ S_2 &= \frac{1}{2} \int_0^s (\sigma_1^2 - 1) \kappa'(\nu_1) \nu_0 ds \end{aligned} \right\} \quad (42)$$

We now apply (41) to determine the modulation fractions f_1 and f_2 of the wanted wave respectively during and after the disturbing square pulse considered in subsection 4.5.

Since $\nu = \nu_0$ at $t=0$ and $M^2(t_0\tau) = 1$ for $0 \leq t \leq T$, we then have $\sigma_1 = 1$ and so (41) yields the formula

$$f_1 = S_1(1 - e^{-t/t_0}). \quad (43.1)$$

Similarly, if $\nu = \nu_s$ at $t=0$ and $M^2(t_0\tau) = 0$ for $t > T$, we have $\sigma_1 = \nu_s/\nu_0$ and so (41) yields

$$f_2 = S_1 e^{-t/t_s}. \quad (43.2)$$

Next we apply (41) to determine the modulation fraction f_i impressed on the wanted wave by the sinusoidally modulated disturbing wave considered in subsection 4.6.

As in 4.6, we now have: $M^2(t_0\tau)$ given by (29.1), $\nu_1 = \nu_s$, i.e.,

$$\sigma_1 = \sigma_s = \nu_s/\nu_0,$$

and (29.2) which is satisfied by the root $\sigma_1 = \sigma_s$.

From (42), (29.2), and (29.1) we obtain

$$S_1 M^2(t_0\tau) + S_2 = S_1 \{ 2M \sin(nt_0\tau) - \frac{1}{2} M^2 \cos(2nt_0\tau) \},$$

and so (41) now becomes

$$\frac{df}{d\tau} + \sigma_1 f = 2S_1 M \sin(nt_0\tau) - \frac{1}{2} S_1 M^2 \cos(2nt_0\tau). \quad (41.1)$$

The impressed modulation fraction is given by the general solution of this equation; namely,

$$f_i = M_1 \sin(nt - \varphi_1) - M_2 \cos(2nt - \varphi_2) + C e^{-t/t_s}, \quad (44)$$

where $\varphi_1, \varphi_2, t_s$ are defined under (33),

$$\left. \begin{aligned} M_1 &= B_s M / (1 + n^2 t_s^2)^{\frac{1}{2}}, \\ M_2 &= \frac{1}{4} B_s M^2 / (1 + 4n^2 t_s^2)^{\frac{1}{2}} \end{aligned} \right\}, \quad (45)$$

$$B_s = -2\sigma_1^{-1} S_1 \quad (46)$$

and C is a constant such that $f_i = 0$ when $t = 0$.

Since $\nu_1 = \nu_s$ we find from (21.1) and (46) that in a region in which ν_0 and E_1 are constant

$$B_s = - \frac{e^2 E_{e1}^2 A_s}{G m^2 l^2 \nu_s} \int_0^{s_w} \kappa'(\nu_s) ds_w \quad (46.1)$$

where A_s is the same as A_1 in (24) with ν_s replacing ν_1 and s_w is the length of the path of the wanted wave.

It will be noted that all the quantities in (44) which relate to the wanted wave are contained in the factor B_s which occurs in the quantities M_1 and M_2 .

A simple calculation shows that when, for the wanted wave, $\kappa(\nu)$ is approximately proportional to ν the formula (44) confirms a previously given formula for M_1 [Bailey, 1937a, part 1, eq (42)].

4.8. Modulation Impressed by a Wave Mode

When it is necessary to take into account the attenuation of each of the wave modes into which a plane wave divides, the theory given above must be modified by using in the place of the formula (13) for \bar{w} a formula for either mode which is obtained as follows [Bailey, 1937a, part 1]:

The mean power Ξ supplied by the wave to the N electrons in unit volume is given by

$$\Xi = -\frac{\partial \bar{P}}{\partial s_d} \quad (47)$$

where \bar{P} is the mean power flux density and s_d is the length of the path of the disturbing wave from its place of entry.

Since this must be equal to N times the mean power $\nu \bar{w}$ supplied to one electron, it follows that

$$N\nu \bar{w} = \Xi. \quad (48)$$

Also, when μ varies slowly with s_d we have from (10) and (9)

$$\Xi = 2\kappa_d \bar{P} = 2\kappa_d \bar{P}_0 \exp(-2\alpha_d) \quad (49)$$

where the subscript d to a symbol relates it to the disturbing wave.

Since the amplitude of the wave mode is modulated by the factor $M(t)$ as in (18), therefore

$$\bar{P}_0 = \bar{P}_{01} M^2(t)$$

where \bar{P}_{01} is the mean flux density of this mode, over one period, at entry into the ionosphere.

Hence, by (49)

$$\Xi = \Xi_1 M^2(t) \quad (50)$$

where

$$\Xi_1 = 2\kappa_d \bar{P}_{01} \exp\left(-2 \int_0^{s_d} \kappa_d ds_d\right) \quad (51)$$

and $\kappa_d \equiv \kappa_d(\nu)$ when expressed as a function of ν .

On using the value of \bar{w} given by (48) and also (50) the equation (19) assumes the form

$$\frac{d\sigma}{d\tau} + \frac{1}{2}(\sigma^2 - 1) = R M^2(t_0 \tau) \quad (52)$$

where

$$\sigma = \nu/\nu_0, \quad \tau = t/t_0 = G\nu_0 t, \quad (53)$$

$$R = \Xi_1 / 2GU_0 N\nu_0 \sigma. \quad (54)$$

As in subsection 4.3 we shall be content here to consider only those situations in which $R M^2(t_0 \tau)$ is small enough for the condition (22.1) to hold true.

Then (52) may be replaced (approximately) by the equation

$$\frac{d\delta}{d\tau} + \sigma_1 \delta = R_1 M^2(t_0 \tau) + \frac{1}{2}(1 - \sigma_1^2) \quad (55)$$

where

$$R_1 = \Xi_1(\nu_1) / 2GU_0 N\nu_1 \quad (56)$$

and $\Xi_1(\nu_1)$ is the value of Ξ_1 in (51) when $\kappa_d = \kappa_d(\nu_1)$ and $\nu_1 \doteq \nu_0$.

Since (55) has the same form as (23) we can now proceed, as was done with the latter, to derive from it the equation for the modulation fraction $f \equiv f_i(t)$ impressed by the wave mode under consideration when the ionospheric region concerned is nearly uniform in temperature and pressure and σ_1 is nearly constant throughout this region.

This procedure yields (41) for f and the formulae (42) but with R_1 now given by (56); i.e., S_1 is now given by the formula

$$S_1 = -\frac{\nu_0}{2GU_0} \int_0^s \Xi_1 q ds \quad (57)$$

where

$$q = \kappa'(\nu_1) / N\nu_1 \quad (58)$$

and Ξ_1 is given approximately by (51) with $\nu = \nu_0$.

We thus see that the formulae (43.1), (43.2), and (44) for the impressed modulations f_1, f_2, f_i , respectively, are also valid here when S_1 is given by (57).

The only factor in S_1 which relates to the disturbing wave is Ξ_1 . For this reason it has been termed the *index of interaction*. Ξ_1 also represents the mean input of power per unit volume of the ionosphere per period of the disturbing wave.

The combined effects of the two modes which constitute a plane wave can be similarly calculated if the distribution of the energy flux between the modes is known.

We now consider the common case (A) in which the wanted wave is an ordinary mode with

$$\kappa(\nu) = \rho N\nu \quad (A)$$

where ρ is nearly constant throughout the region concerned. In this case we generally have

$$\omega^2 \gg \nu^2 \text{ and } \omega^2 \gg 4\pi N e^2 / m.$$

Therefore, by (58), $q = \rho/\nu_1$ and so q is now nearly constant throughout the region. Hence, by (57) we now have

$$S_1 \doteq -\frac{\kappa(\nu_0)}{2GU_0 N\nu_1} \int_0^s \Xi_1 ds. \quad (57.1)$$

Since, by (50) and (47),

$$M^2(t) \int_0^s \Xi_1 ds = - \int_0^{s_d} \sec \beta \frac{\partial \bar{P}}{\partial s_d} ds_d \quad (59)$$

where β is the angle between the elements of path ds_d and ds contained between adjacent wave fronts of the disturbing wave,¹⁰ we cannot evaluate the integral in (57.1) in finite terms except when β is constant (or nearly so) throughout the region of interaction; this will occur in the following subcases, A1 and A2.

A1. Both waves are propagated vertically and then $|\sec \beta|=1$ throughout.

A2. The disturbing wave is propagated vertically and the region of interaction occurs mainly where the wanted wave is undeviated.

In these cases (59) yields

$$M^2(t) \int_0^s \Xi_1 ds = \sec \beta (\bar{P}_0 - \bar{P}_{s_d})$$

where \bar{P}_0 and \bar{P}_{s_d} are the values of the disturbing wave's flux density at the beginning and end of a length s_d of its path.

Since

$$\bar{P}_0 = \bar{P}_{01} M^2(t) \text{ and } \bar{P}_{s_d} = \bar{P}_{s_d1} M^2(t)$$

where \bar{P}_{01} and \bar{P}_{s_d1} are the mean flux densities over one period at, respectively, entry into and exit from the ionosphere, it follows that

$$\int_0^s \Xi_1 ds = \sec \beta (\bar{P}_{01} - \bar{P}_{s_d1}),$$

and so from (57.1) that, in the subcases A1 and A2,

$$S_1 = -\frac{\kappa(\nu_0) \sec \beta}{2GU_0 N \nu_1} (\bar{P}_{01} - \bar{P}_{s_d1}). \quad (60)$$

Under A2 we have $\sec \beta = \sec i_w$ where i_w is the angle of incidence of the wanted wave. We then obtain from (44) and (60) a formula for the depth M_1 , of the impressed fundamental modulation, which is equivalent to the equation (28) given in Huxley's Synopsis [1952] and is originally due to Ratcliffe and Shaw [1948].

4.9. Gyro Interaction: Interaction Caused by an Approximate Gyrowave

When the disturbing wave is plane-polarized and has a frequency within 20 percent of the local gyro-frequency, i.e.,

$$\left. \begin{aligned} \omega &= \Omega(1 + \epsilon) \\ |\epsilon| &\leq 0.2 \end{aligned} \right\}, \quad (61)$$

the formula (24) with $\nu_1 = \nu_s$ shows that in a region where

$$\nu_s^2 < \Omega^2 \quad (62)$$

the quantity A_s passes through a sharp maximum as the wave's frequency ω is varied about the gyro-frequency Ω .

It follows from (46.1) that B_s behaves similarly, and that in consequence each of the modulation depths M_1 and M_2 passes through a sharp maximum. By (21.1), (42), and (43.1), the modulation f_1 impressed by a pulsed gyrowave will behave similarly.

Since the two component modes of the plane gyrowave are absorbed at very different rates as they propagate through the ionosphere, a more precise discussion requires the effects caused by each of these modes to be considered separately. We may then use the results obtained in subsection 4.8.

For our present purpose it is sufficient to consider only the effects caused by the much more absorbable extraordinary component¹¹ in the lower E region of the nocturnal ionosphere, at moderate latitudes, for which we have

$$N < 800, \Omega = 10^7, \nu \leq 10^6. \quad (63)$$

It has been shown [Bailey, 1938a] that the refractive index μ of the extraordinary mode defined by (61) is closely equal to unity and that its coefficient of absorption κ_d is given approximately by

$$\kappa_d = \frac{K(1 + \cos^2 \theta) N \nu}{\nu^2 + \epsilon^2 \Omega^2} \quad (64)$$

where

$$K = 0.0266,$$

and θ is the angle between the directions of propagation and of the earth's magnetic field.

More recently R. A. Smith [1957] has obtained the following closer approximation:

When

$$N < 500, \Omega = 10^7, 3 \times 10^5 < \nu < 1.2 \times 10^6, |\theta| < 65^\circ,$$

then

$$\kappa_d = \frac{K(1 + \cos^2 \theta) N \nu}{\nu^2 + \epsilon^2 \Omega^2 - p_0^2 (A_1 \epsilon - A_2 \epsilon^2)} \quad (65)$$

where

$$\left. \begin{aligned} p_0^2 &= 3.19 \times 10^9 N \\ A_1 &= \frac{1}{4}(2 + 3 \sin^2 \theta) \\ A_2 &= \frac{1}{2}(1 + 3 \sin^2 \theta) \end{aligned} \right\}. \quad (66)$$

From (64) or (65) it is clear that under the specified conditions κ_d passes through a notable maximum value as ω varies about the value Ω . On the other hand (51) shows that Ξ_1 may then pass through either one or two maxima according as the path length s_d is shorter or longer than some particular value. As shown by (57) this behavior will be reflected in S_1 through the integral

$$\int_0^s \Xi_1 q ds,$$

¹¹ When required, the weaker effects due to the ordinary component can be calculated in a similar way.

¹⁰ For which $ds_d = ds \cos \beta$.

and therefore, by (43.1) and (44), the modulations f impressed by a pulsed or sinusoidally modulated gyro-wave will pass through one or two maxima as the disturbing wave's frequency varies about the gyro-frequency. The curves representing S_1 , M_1 , or M_2 in terms of ϵ , ω , or $\omega/2\pi$, may therefore be called resonance curves.

Whether the extraordinary gyrowave is transmitted from the ground or from a rocket or a satellite in (or above) the ionosphere, calculation shows that in general the wave is almost completely absorbed within a region of the ionosphere about 6 km long.

Consequently we may expect the interaction to be greatest when the ionospheric path of the wanted wave lies within this region.

With the gyrowave transmitted from the ground this region at night lies in a horizontal slab with its lower face at about 90 km and, for $\epsilon=0$, its upper face at about 95 km above the ground. As $|\epsilon|$ increases from 0 to 0.2 the upper face rises from about 95 km to about 97 km.

Consequently, when the top of the wanted wave's path lies below 95 km, the resonance curve for the impressed modulation in terms of ϵ possesses only one maximum (dromedarian resonance). But when this top lies above 95 km the resonance curve possesses two maxima nearly symmetrically situated about the minimum which occurs near $\epsilon=0$ (bactrian resonance), for increase of $|\epsilon|$ from 0 initially causes the slab to embrace an increasing portion of the wanted wave's path and so increases the interaction.

This deduction is supported by the published curves for the index of interaction Ξ_1 at different levels in a specified model of the lower ionosphere [Bailey, 1938a]. For these have one maximum at the lower levels and two maxima at the higher levels.

For a particular model of the nocturnal ionosphere and the conditions relating to experiments on gyro interaction at Armidale, New South Wales, during the period 1950–1954, R. A. Smith [1957] has carried out the computations which are needed to evaluate qds and the integral in (57). These computations led to a theoretical resonance curve for the impressed modulation which was in agreement with the observed resonance curves [Bailey, Smith, et al., 1952] within the limits of experimental error. Some of these curves are shown in figure 3 below.

When the disturbing and wanted waves are both propagated vertically, the computation of the integral in (57) is notably simplified. The experimental work is also much simplified provided that the power is increased sufficiently to compensate for the reduction in interaction which would otherwise occur.

If also the wanted wave is an ordinary mode under the conditions discussed above, (57) may be replaced by (57.1) and so the theoretical computations then become simpler still.

The impressed modulations predetermined by means of this theory of gyro interaction have been fully confirmed within the limits of experimental error by subsequent experiments.

5. Predictions From the Mathematical Theory

The formulae (44), (45), and (46.1) lead to the predictions labeled (1) in section 1.

The formulae (45) specify the manners in which M_1 and M_2 decrease with increasing n , as stated in the prediction labeled (2). These variations with the modulation frequency are similar to the variation of the current in a choke coil as the frequency of the applied p.d. is varied.

The formula (46.1) shows that when E_{e1}^2 is not too large, B_s is nearly proportional to E_{e1}^2 and, therefore, also to the mean power \bar{P} of the disturbing station. It then follows from (45) that M_1 and M_2 are nearly proportional to \bar{P} and so we obtain the prediction (3).

The formulae (44) and (33) show that

$$\varphi_1 = \tan^{-1}(n/G\nu_s), \quad \varphi_2 = \tan^{-1}(2n/G\nu_s).$$

Hence, when the mean power of the disturbing wave is such that ν_s , the steady value of the perturbed collision frequency, differs only slightly from ν_0 , the unperturbed value, we obtain the prediction (4).

The formulae (45) and (46.1) show that M_1 and M_2 are proportional to A_s , where A_s is given by (24) with ν_1 replaced by ν_s . Hence, when ν_s differs only slightly from ν_0 , we obtain the predicted behavior (5) and also the single maxima predicted under the label (6). But the more precise discussion of subsection 4.9 shows that when s_d the path of the disturbing wave exceeds a certain value each of the modulations depths passes through two maxima. Also, computations by means of the formulae given in subsection 4.9 show that when $\epsilon=0$ a mean power of 1 or 2 kW radiated from a vertical aerial suffices to produce observable interaction. All this is summed up in the prediction (6).

The rays from an average radio station which produce the largest values of E_{e1}^2 in the E -region are inclined at 45° to the vertical. Consequently the ionospheric regions where a wanted wave is most modulated by the disturbing station will be about 100 km from the ionospheric region vertically above this station. This is the prediction (7).

The formulae (43.1), (43.2), (16), and (33) show that, when the disturbing wave is such that ν does not differ much from ν_0 , the modulation impressed by a pulsed disturbing wave rises and falls approximately according to the same exponential law with the time constant $t_0=1/G\nu_0$. This yields the prediction (8).

As the mean power \bar{P} of a wave increases the modulations of its amplitude with the same depth M , but different frequencies n will, by (32), cause modulations of ν of depths which are greatest for the lowest frequencies. Consequently the modulations of the lowest frequencies will be received with the smallest amplitudes. This self-distortion must increase with \bar{P} and so we obtain the pre-

diction (9). Quantitative discussions of self-distortion have been given by Hibberd [1955, 1957].

The prediction (10) follows immediately from the fact that the low-frequency radiation from an electric storm reaches the nocturnal *E*-region with great intensity and so considerably increases the local electron collision frequency for periods of the order of one millisecond.

The same general mathematical theory [Bailey, 1937b] which yielded (11), for a single disturbing wave, also yielded a formula [there labeled (35)] for the variation of the collision frequency caused by two or more disturbing waves. The latter formula contained terms representing combination tones and so the prediction (11) immediately follows.

The predicted phenomena (12) and (13) are of great interest and possible future importance for the study of the ionosphere and ionospheric radio propagation. Since no experimental tests of the prediction (12) have as yet been carried out, it is sufficient here to refer to the original publications in which it is discussed in detail [Bailey, 1938, 1959]. The predictions (13) are under experimental test by Professor R. A. Smith at the University of New England in New South Wales, as indicated in section 8.

6. Experimental Tests of the Theory and Verification of the Predictions

The principal experiments which served to test one or more of these predictions will now be mentioned in their historical order.

Van der Pol and van der Mark [1935] measured the depth of modulation M impressed on Radio Beromunster by Radio Luxembourg at four different frequencies of modulation n . Their observations are represented by the four dots shown in figure 1. The theoretical curve representing M_1 as a function of n , with appropriately selected values of the two constants ($B_s M$) and t_s in (45), is also shown there. The resulting value of $G\nu_0 \div 1/t_s$ then yielded a value of ν_0 which was consistent with the values of ν_0 previously estimated by E. V. Appleton [1932] and S. Chapman [1932] from other experiments. This constituted the earliest verification of the prediction (2) with regard to M_1 . Also the measured depths of modulation were later found to agree in order of magnitude with the theoretical value M_1 given in (45), thus verifying a part of the prediction (1).

Under the guidance of E. V. Appleton [1934] and Ralph Stranger [1934], the members of the World Radio Research League in Europe made observations which roughly verified the prediction (7).

Bäumler and Pfitzer [1935] qualitatively verified predictions (2), (5), and (7).

Extensive experiments which were carried out in 1937 with the help of the British Broadcasting Corporation and other European authorities verified at least the last part of the prediction (6) [Bailey, 1937d]. During these experiments, E. W. B. Gill of Merton College, Oxford, reported observations which appeared to verify prediction (11).

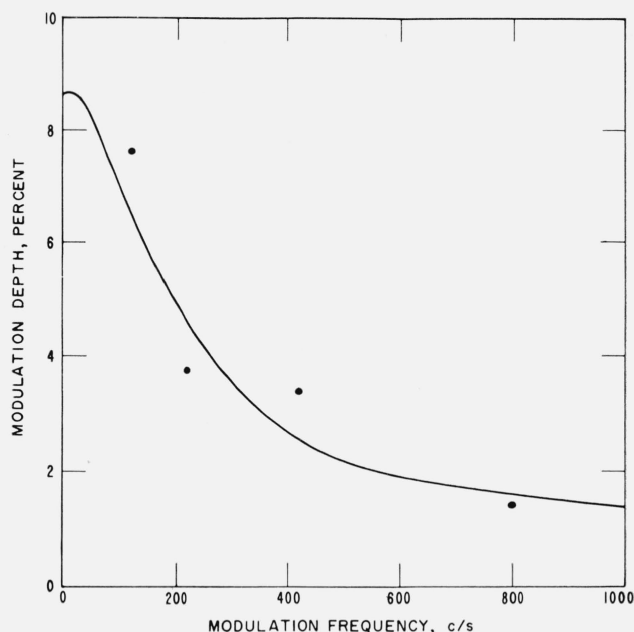


FIGURE 1. Modulation frequency, c/s.

[Bailey, V. A., and D. F. Martyn, *Nature* **135**, 585 (1935).]

From 1946 onwards M. Cutolo of Naples and his colleagues successfully made extensive observations of gyro interaction using disturbing waves with powers less than 1 kW. They also succeeded in demonstrating the existence of both double-humped (bactrian) and single-humped (dromedarian) resonance. Thus they verified the whole of the prediction (6).

Commencing in 1946 L. G. H. Huxley and J. A. Ratcliffe and their respective colleagues carried out extensive experiments, both separately and jointly, which verified the predictions (1), (2), (3), and (4) (excepting the proportionality of M_2 to M^2). Moreover, Huxley, Foster and Newton [1947] observed the gyro interaction caused by the 10 kW station at Stagshaw; this is consistent with the last part of prediction (6). In addition, Ratcliffe and Shaw [1948] directly verified the fundamental concept of the theory that the amplitude of the wanted wave is decreased when the disturbing wave is operating. Also Ratcliffe and Shaw successfully applied the theoretical formula for the phase lag ϕ_1 , given under (33), to determine experimentally the height of the ionospheric region at which the wave interaction took place. Since the heights so determined were of the same order as the heights postulated in the original theory, we may regard this work as an experimental verification of the prediction (4).

In 1949 experiments on gyro interaction were commenced in Eastern Australia by V. A. Bailey and his colleagues at the University of Sydney in collaboration with R. A. Smith and his colleagues at the University of New England at Armidale (N.S.W.). These were planned to conform as closely as possible to the published requirements of the theory [Bailey, 1937c and 1938a]. Accordingly

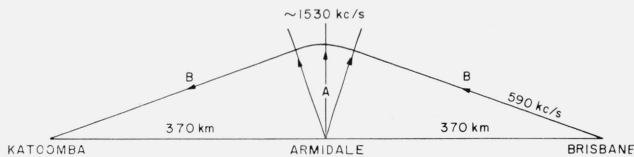


FIGURE 2.

the arrangement of stations shown in figure 2 was adopted. In this, the disturbing gyrowave *A* was radiated vertically from a horizontal aerial at Armidale, which lies at the midpoint of the 740 km line that joins Brisbane and Katoomba. The wanted wave *B* was radiated from Brisbane on 590 kc/s and received at Katoomba. The local gyrofrequency at Armidale was estimated to be 1530 kc/s.

The gyrowave was radiated 40 times per second as a square pulse of length 1 msec and power 36 kW. Its frequency f was kept constant for 2 min at each of 8 values lying within ± 23 percent of the gyrofrequency 1530 kc/s. Each of the cycles 1, 2, and 3 ran through the middle 6 frequencies and lasted 17 min. Cycle 4 ran through all the 8 frequencies and lasted 23 min.

In order to provide a standard of comparison for the depth of the modulation impressed on the wanted wave *B*, the latter was modulated at its source for 30 sec in each 2 min period with a standard tone of 80 c/s to a depth of 5 percent. This depth was chosen because computations based on our mathematical theory of gyro interaction had indicated in advance that the experimental arrangement shown in figure 2 would impress on *B* modulation depths of the order of 5 percent. This prediction was nicely verified by the experimental observations.

A detailed account of these Australian experiments on gyro interaction was published by Bailey, Smith, Landecker, Higgs, and Hibberd [1952]. A briefer account was published by Bailey in 1956. Examples of the experimental resonance curves are shown in figure 3. It will thus be seen that these experiments not only verified the predictions (6) and (8), but also confirmed the theory quantitatively in every particular.

The theory of interaction, including gyro interaction, was also thoroughly verified by the experiments on microwave interaction in gas-discharge tubes of Goldstein, Anderson, and Clarke [1953] and of Anderson and Goldstein [1955, 1956]. In some of these experiments Goldstein was able to demonstrate also the self-distortion of a powerful wave [Private communication, 1955], thus verifying the prediction (9) for a plasma.

The prediction (9), that self-distortion in the ionosphere increases with the power and does so at the expense of the lower modulation frequencies, was verified by J. W. King [1957].¹² In carrying

out his experiments he was guided by the quantitative theory of self-distortion developed by F. H. Hibberd [1955, 1956].

Other tests of the theory arose from the experimental study of wave interaction by J. A. Fejer [1955] in which both the wanted and disturbing waves were pulse-modulated. This study demonstrated the existence of interaction during the day at heights between 70 and 90 km and also that a modification of the theory suggested by Huxley was not valid.

The prediction (5) is now verified by the fact that during the last 30 years the greatest disturbance by interaction backgrounds has been produced by the radio stations with the lowest wave frequencies.

From the whole discussion given above we see that experimental observations have verified the 10 predictions (1) to (9) and (11). The predictions (10), (12), and (13) remain to be tested. We may therefore conclude that in the form presented here the theory of wave interaction is experimentally established.

7. Some Applications to Research on the Ionosphere and Gas Discharges and to Radio Communication

The experiments mentioned above by van der Pol and van der Mark; Cutolo, Huxley, Ratcliffe, and Shaw; Bailey and Smith; and Fejer, combined with the theory, have among them served to measure the following important quantities relating to the ionosphere:

The electron collision frequencies ν , the electron densities N , the local magnetic fields H , the constant G (relating to the mean loss of energy by an electron at a collision with a gas molecule), and the heights at which the wave interaction occurs.

Similar applications to the *D*-region, have been made more recently by Fejer and Vice [1959], Bjelland et al. [1959], and Barrington and Thrane [1962].

The experiments by Goldstein and his associates, mentioned above, also demonstrate the value of the phenomenon of wave interaction and its theory for investigating several fundamental processes in a gas discharge.

As an illustration of its practical value for radio communication, it may be mentioned that the author has been consulted about it more than once by the leading broadcasting organizations in Europe and the United States. Also, in the URSI Information Bulletin No. 73, on p. 59, it was reported that the C.C.I.R. stressed the practical importance of information on nonlinear effects in the ionosphere.

The present trend towards the use of increasingly powerful beams in connection with the reception of irregularly scattered waves has also compelled attention to the possible occurrence of observable wave interaction and other nonlinear phenomena at very high frequencies.

¹² In this publication King inadvertently contradicts the statement on "Ionospheric self-interaction" previously made by Hibberd [1955]; namely: "Its existence was first predicted by V. A. Bailey in 1935, . . ."

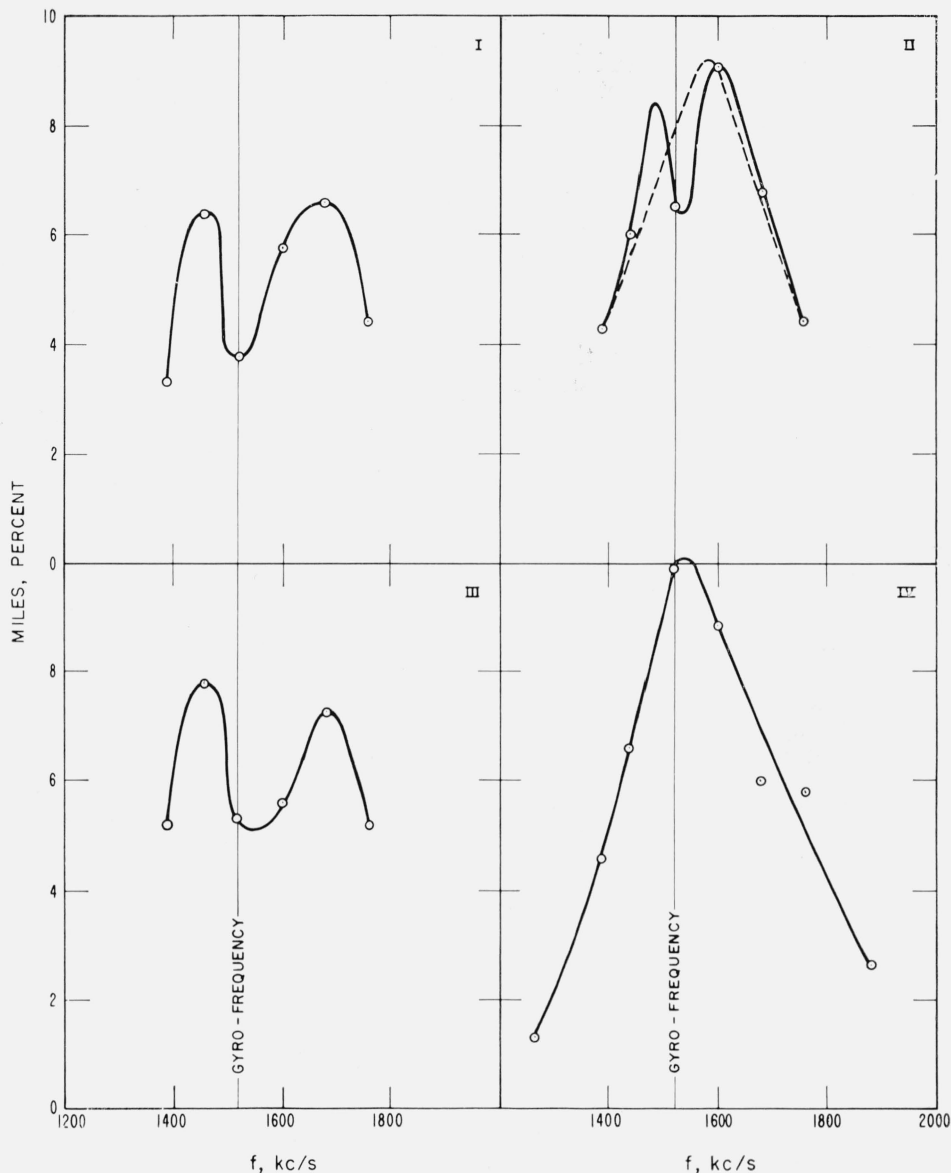


FIGURE 3. [Reproduced from Bailey, V. A., R. A. Smith, K. Landecker, A. J. Higgs, and F. H. Hibberd, *Nature* **169**, 912 (1952).]

Lastly, the possibility mentioned by Bailey and Goldstein [1958] of controlling the ionosphere in various ways by means of powerful radio waves has attracted some attention.

8. Some Recent Developments

Filling in the gap between the two situations corresponding to wave-interaction caused by moderate powers and ionospheric airglows caused by extremely high powers and large aerial arrays, an investigation was made [Bailey, 1959] of the possible increases of both the collision frequencies ν and the electron densities N caused by a powerful extraordinary circular gyrowave in the nocturnal

lower E -region and in the daytime D -region. The theory was based on the most reliable available experimental data on the behavior of free electrons in air with energies up to 1.22 eV. The principal process then freshly introduced was the attachment of electrons to molecules in some collisions.

It was found that some of the resultant effects, of such increases in ν and N , on this wave or on other waves are remarkable in magnitude or in kind. Thus, when the gyrowave produces near the 92 km level at night a reduced equivalent electric force (Z/p) equal to 2.5 V/cm per mm Hg (measured at 288 °K), it causes notable selfdistortion and extremely large wave interaction; also within 3 msec its penetration into the E -region increases considerably.

Other possible observable effects which may be caused at night when Z/p has a value between 2.5 and 8 are changes in the visible and infrared airglow, in parts of some trails of meteors and fast particles of solar or cosmic origin, in the radar echoes from such trails and in the local magnetic elements.

The theory also showed that a pulse of gyro-waves such as that specified above could be used to determine by experiments on the ground the rates of attachment of electrons and other electron-collision data pertaining to the regions near the 92 and 75 km levels.

For these and other reasons, it was therefore suggested that experiments be carried out with gyrowave pulses beamed from an array of at least 80 horizontal dipoles connected to a generator of at least 500 kW power.

As a result of this work and of my recommendation to the Air Force Cambridge Research Center at Bedford (Mass.), the University of New England is under contract with this research center to carry out a project to "Modify the Ionosphere by means of Gyrowaves" using an array of 40 dipoles connected to a generator of 500 kW power. The results of this project are being reported by R. A. Smith.

The power density requirements for the excitation of an artificial airglow by means of gyrowaves [Bailey, 1938b] have been reconsidered [Bailey, 1961] in the light of later information. One important conclusion was that an array of 80 dipoles radiating vertically a beam of gyrowaves with a mean power of 500 kW would produce an easily observable enhancement of the nocturnal airglow overhead within a solid angle of 32 square degrees. Another conclusion is that from the same array pulses of power 3500 kW and length about 0.2 msec would each produce a flash of airglow that is 40 times as bright as the night sky.

An extensive report on nonlinear phenomena in a plasma was published by Ginzburg and Gurevich in 1960 and is available in an English translation [1960 a & b]. The theory is based on the Boltzmann equation for the energy distribution function of the electrons, thus following the classical examples of H. A. Lorentz and F. B. Pidduck. The authors point out that "to disclose the physical picture" the simpler "elementary theory" is "convenient and useful." Our present account of the "elementary theory" confirms this opinion.

Similar discussions based on Boltzmann equations for the electronic distribution functions are contained in reports by Megill [1961], Carleton and Megill [1962], and Molmud et al. [1962]. These are directed towards special problems of considerable interest such as, respectively, the generation of artificial airglows and a method to reduce the electron density in the D-region, by means of powerful radio transmitters on the ground. In all these discussions the use of simplifying approximations was found necessary.

An interesting new way of using wave interaction for studying the ionosphere has been proposed inde-

pendently by Vilenski [1962] and Ferraro, Lee, and Weisbrod [1963]. This depends on measurement of the phase changes in the wanted wave that are caused by the disturbing wave.¹³

A remarkable new example of the interaction of combination tones of two high-frequency waves with a third (wanted) wave, similar to that mentioned under prediction (11) in sections 1 and 5, has been reported by Cutolo [1962].

The waves concerned and our theoretical processes were as follows:

(1) Disturbing wave: A carrier wave of frequency $\omega/2\pi$ between 50 and 75 Mc/s was modulated at a frequency of $\Omega/2\pi$ between 1000 and 1300 kc/s which lies near the local gyrofrequency of 1200 kc/s; the peak power was 80 kw and the modulation depth m about 42 percent. Thus the (Yagi) aerial used simultaneously radiated three waves of the angular frequencies $\omega + \Omega$, ω , and $\omega - \Omega$. Two pairs of these produced in the ionosphere electric currents which contained components with frequencies equal to the combination tones $(\omega + \Omega) - \omega$ and $\omega - (\omega - \Omega)$, i.e., equal to the gyrofrequency Ω . These ionospheric current components generated secondary waves of the frequency Ω , which in turn caused gyroresonance variations of the local collision frequency ν .

(2) The wanted waves were selected so as to be reflected at or above the 90 km level in the ionosphere in accordance with the requirements for gyro interaction given in our section 4.9.

Cutolo's experiments demonstrated that in this way modulations of depths from 5 to 20 percent could be impressed on the wanted waves near the 90 km level of the ionosphere when the modulation frequency approximated to the gyrofrequency.

Lastly, we may consider the waves that could be generated in the interplanetary plasma by the large negative charge on the sun, which has been postulated in recent years to account simultaneously for a number of important phenomena [Bailey, 1960]. The evidence in support of this hypothesis has recently been strengthened by the fact that three predictions (based on this hypothesis) that were published in 1960 have now been verified by the data on the interplanetary magnetic field obtained by means of the space probes Pioneer 5, Explorer 10, Mariner 2, and Explorer 12 [Bailey, 1963].

The electrostatic field E_s due to the charge $-Q_s$ on the sun is radial to the sun and has the magnitude

$$E_s = fQ_s/r^2 \text{ e.s.u.}$$

at the radial distance r , where f is a number between 0 and 1 which represents the partial screening of the solar charge by the interplanetary plasma clouds.

Any fluctuations in the emission by the sun of these plasma clouds will cause corresponding fluctuations E in the field E_s . In certain frequency bands the radial components of E are particularly liable to

¹³ Ferraro and his coauthors mistakenly attribute the origin of the theory of wave interaction to its discussion in 1948 instead of to the correct date 1934.

amplification by the mean static field E_s , since they may be regarded as forming longitudinal wave modes.

These longitudinal waves may be sufficiently strong to produce, or help produce, certain notable nonlinear effects in the earth's ionosphere, which are at present attributed to other causes. For example, some geomagnetic disturbances¹⁴ may be caused by such longitudinal waves [Heirtzler, 1962].

This possibility deserves further study, both by means of instrumented satellites and theoretically [Bailey, 1948, 1951].

9. Conclusion

The foregoing account of the theory of wave interaction and other nonlinear processes in the ionosphere in the form developed at the University of Sydney shows how the ideas and methods of the Townsend School on the motions of electrons in gases and of the Appleton School on magneto-ionic wave propagation have led to a relatively simple mathematical theory which has proved fertile in predictions of new phenomena that have subsequently been verified by experiment.

It may therefore be used, and has been used, with confidence as the basis of new experimental methods for investigating processes in the ionosphere and in electrical gas discharges. It has also suggested methods of controlling parts of the ionosphere for particular purposes by means of beams of radio waves.

The present trend towards increasingly large power densities in radio beams has made this theory, and its mathematical elaborations, of increasing practical importance.

These elaborations, which are based in part on Boltzmann's equation for the distribution of electron energies, are distinctly onerous, but they serve both to check the simple mathematical theory of the Townsend School and to study *ab initio* those phenomena which, like the artificial airglow, depend strongly on the actual distribution of the electron energies. On account of continual and nonuniform fluctuations of the ionosphere and the lack of precision in the necessary auxiliary laboratory data, it is probable that for some time to come the simple theory¹⁵ will serve for most purposes quite as well as the elaborations.

The present account also suggests that certain types of low-frequency longitudinal waves of extraterrestrial origin may produce some modifications of the ionosphere that are at present attributed to other causes.

10. References

- Anderson, J. M., and L. Goldstein (1955), Interaction of electromagnetic waves of radio-frequency in isothermal plasma: collision cross-section of helium atoms and ions for electrons, *Phys. Rev.* **100**, No. 4, 1037-1046.
- Anderson, J. M., and L. Goldstein (1956), Momentum transfer cross section and fractional energy loss in the collisions of slow electrons with nitrogen molecules, *Phys. Rev.* **102**, No. 2, 388-389.
- Appleton, E. V. (1932), Wireless studies of the ionosphere, *J. IEE* **71**, 642-650.
- Appleton, E. V. (Nov. 30, 1934), *World Radio*. Title unavailable.
- Appleton, E. V. (Dec. 7, 1934), *World Radio*. Title unavailable.
- Bailey, V. A. (Oct. 1925), On the attachment of electrons to gas molecules, *Phil. Mag.* **50**, No. 298, 825-843.
- Bailey, V. A. (Jan. 9, 1937a), Interaction by resonance of radio waves, *Nature* **139**, 68-69.
- Bailey, V. A. (April 1937b), The motions of electrons in a gas in the presence of variable electric fields and a constant magnetic field, *Phil. Mag.* **23**, No. 156, 774-791.
- Bailey, V. A. (May 1937c), On some effects caused in the ionosphere by electric waves. Part I, *Phil. Mag.* **23**, No. 157, 929-960.
- Bailey, V. A. (May 15, 1937d), Resonance in the interaction of radio waves, *Nature* **139**, 838-839.
- Bailey, V. A. (Oct. 1938a), On some effects caused in the ionosphere by electric waves. Part II, *Phil. Mag.* **26**, No. 176, 425-543.
- Bailey, V. A. (Oct. 1, 1938b), Generations of auroras by means of radio waves, *Nature* **142**, 613-614.
- Bailey, V. A. (1948), Plane waves in an ionized gas with static electric and magnetic fields, *present Australian J. Sci. Res.* **1**, 351-359.
- Bailey, V. A. (July 15, 1951), The relativistic theory of electro-magnetoionic waves, *Phys. Rev.* **83**, No. 2, 439-454.
- Bailey, V. A. (1956), On the interaction of radio waves, *Il Nuovo Cimento* **4**, (Suppl. 4), 1430-1449.
- Bailey, V. A. (1959), Some possible effects caused by strong gyro-waves in the ionosphere—I., *J. Atmospheric Terrest. Phys.* **14**, 299-324.
- Bailey, V. A. (1960), *Proc. Roy. Soc. N.S. Wales* **94**, 77. Title unavailable.
- Bailey, V. A. (1961), Power density requirements for airglow excitation by gyrowaves, *J. Res. NBS* **65D**, (Radio Prop.), No. 4, 321-322.
- Bailey, V. A. (Sept. 14, 1963), The interplanetary magnetic field, *Nature* **199**, 1029-1031.
- Bailey, V. A., and D. F. Martyn (Feb. 10, 1934), Interaction of radio waves, *Nature* **133**, 218.
- Bailey, V. A., and D. F. Martyn (Apr. 13, 1935), Interaction of radio waves, *Nature* **135**, 585.
- Bailey, V. A., R. A. Smith, K. Landecker, A. J. Higgs, and F. H. Hibbert (May 31, 1952), Resonance in gyro-interaction of radio waves, *Nature* **169**, 911-913.
- Bailey, V. A., and L. Goldstein (1958), Control of the ionosphere by means of radio waves, *J. Atmospheric Terrest. Phys.* **12**, 216-217.
- Barrington, R. E. and E. Thrane (1962), The determination of D-region electron densities from observations of cross modulation, *J. Atmospheric Terrest. Phys.* **24**, 31-42.
- Bäumler, M., and W. Pfitzer (1935), *Hochfreq. Elektroak.* **46**, 81. Title unavailable.
- Bjelland, B., O. Holt, B. Landmark, and F. Lied (Sept. 26, 1959), The D-region of the ionosphere, *Nature* **184**, 973-974.
- Butt, A. G. (Apr. 28, 1933), *World Radio*. Title unavailable.
- Carleton, N. P., and L. R. Megill (1962), Electron energy distribution in slightly ionized air under the influence of electric and magnetic fields, *Phys. Rev.* **126**, 2089-2099.
- Chapman, S. (1932), The lunar diurnal variation of atmospheric temperature at Batavia, 1866-1928, *Proc. Roy. Soc.* **A137**, 1-24.
- Crompton, R. W., L. G. H. Huxley, and D. J. Sutton (1953), Experimental studies of the motions of slow electrons in air with application to the ionosphere, *Proc. Roy. Soc.* **A218**, 507-519.
- Cutolo, M. (Dec. 13, 1947), Gyro-interaction of radio-waves obtained by the pulse method, *Nature* **160**, 834.
- Cutolo, M. (July 15, 1950), Effects of radio gyrointeraction and their interpretation, *Nature* **166**, 98-100.
- Cutolo, M. (1952), Determinazione sperimentale delle curve di risonanza nel moto degli elettroni lenti dell'alta atmosfera, *Il Nuovo Cimento* **9**, 391-406.

¹⁴ For example the micropulsations of shorter type A waves (period ~ 1 sec) modulated by type C waves of periods about 20 sec [Sci. Am., **206**, 128 (Mar. 1962)].

¹⁵ Also called, by Ginsburg, the "elementary theory."

- Cutolo, M. (June 1962), Istituto di Fisica Tecnica, University of Naples. Title unavailable.
- Cutolo, M., M. Carlevaro, and M. Gherghi (1946), Esperienze sull'interazione con risonanza fra radioonde nella ionosfera, *Alta Frequenza* **15**, No. 2, 111-117.
- Fejer, J. A. (1955), The interaction of pulsed radio waves in the ionosphere, *J. Atmospheric Terrest. Phys.* **7**, 322-332.
- Fejer, J. A., and R. W. Vice (1959), An investigation of the ionospheric *D*-region, *J. Atmospheric Terrest. Phys.* **16**, 291-306.
- Ferraro, A. J., H. S. Lee, and S. Weisbrod (1963), Phase interaction: a new tool for *D*-region studies, *J. Geophys. Res.* **68**, No. 4, 1169-1171.
- Gill, E. W. B. (Mar. 1935), Private communication.
- Ginzburg, V. L., and A. V. Gurevich (1960a), Nonlinear phenomena in a plasma located in an alternating electromagnetic field, *Soviet Physics, Uspekhi* **3**, 115-146.
- Ginzburg, V. L., and A. V. Gurevich (1960b), Nonlinear phenomena in a plasma located in an alternating electromagnetic field, *Soviet Physics, Uspekhi* **3**, 175-194.
- Goldstein, L., J. M. Anderson, and G. L. Clark (1953a), Interaction of microwaves propagated through a gaseous discharge plasma. *Phys. Rev.* **90**, No. 1, 151-152.
- Goldstein, L., J. M. Anderson, and G. L. Clark (1953b), Quenching of afterglow in gaseous discharge plasma by low power microwaves, *Phys. Rev.* **90**, No. 1, 486-487.
- Healey, R. H., and J. W. Reed (1940), The Behaviour of Slow Electrons in Gases (A. W. A. Ltd. Sydney, Australia).
- Heitzler, James R. (March 1962), The longest electromagnetic waves, *Sci. Am.* **206**, 128-137.
- Hibberd, F. H. (1955), Ionospheric self-interaction of radio waves, *J. Atmospheric Terrest. Phys.* **6**, 268-279.
- Hibberd, F. H. (1956), On the observation of ionospheric self-interaction, *J. Atmospheric Terrest. Phys.* **8**, 120-122.
- Hibberd, F. H. (1957), Self-distortion of radio waves in the ionosphere, near the gyro frequency, *J. Atmospheric Terrest. Phys.* **11**, 102-110.
- Huxley, L. G. H. (1950), Ionospheric cross-modulation at oblique incidence, *Proc. Roy. Soc.* **A200**, 486-511.
- Huxley, L. G. H. (1952), A synopsis of ionospheric cross-modulation, *Il Nuovo Cimento* **9** (Suppl. Series 9), 59-89.
- Huxley, L. G. H., H. G. Foster, and C. C. Newton (Mar. 1, 1947), Gyro interaction of radio waves, *Nature* **159** 300-301.
- Huxley, L. G. H., and J. A. Ratcliffe (1950), Discussion on "A survey of ionospheric cross-modulation," *Proc. IEE* **97**, Part III, 165.
- King, J. W. (1957), Ionospheric self-demodulation of radio waves, *J. Atmospheric Terrest. Phys.* **10**, 166-167.
- Llewellyn Jones, F. (1953), Electrical discharges, Reports on progress, *Physics* **16**, 216.
- Megill, L. R. (1961), Expected upper atmospheric effects in the presence of high frequency fields with high power density (private communications).
- Molmud, P., S. Altshuler, and J. H. Gardner (September 1962), The Study of a Method to Reduce Electron Density in the Ionospheric *D*-region by Means of High-powered Ground-based Transmitters. Space and Technology Labs., Inc., Final Report.
- Nielsen, Russel A., and Norris E. Bradbury (Jan. 15, 1937), Electron and negative ion mobilities in oxygen, air, nitrous oxide and ammonia, *Phys. Rev.* **51**, 69-75.
- Ratcliffe, J. A. (1948), The interaction of radio waves, *J. IEE* **95**, Part III, No. 33, 325.
- Ratcliffe, J. A., and I. Shaw (1948), A study of the interaction of radio waves, *Proc. Roy. Soc.* **A193**, 311-343.
- Smith, R. A. (1957), Doctoral Thesis, University of Sydney, Sydney, Australia.
- Stranger, Ralph (Dec. 14, 1934), *World Radio*.
- Tellegen, B. D. H. (June 10, 1933), Interaction between radio-waves?, *Nature* **131**, 840.
- Townsend, J. S. (Apr. 1932), Electrodeless discharges, *Phil. Mag.* **S-7**, **13**, No. 86, 745-759.
- Townsend, J. S. (1947), *Electrons in Gases* (Hutchinson's Publications).
- Townsend, J. S., and H. T. Tizard (1913), The motion of electrons in gases, *Proc. Roy. Soc.* **A88**, 336-347.
- Van der Pol, B. (1935), *Tijdschrift Nederland, Radio-genootschap* **7**, 93.
- Van der Pol, B., and J. van der Mark (1935), *Tijdschrift Nederland, Radiogenootschap* **7**, 12.
- Vilenski, I. M. (1962), *Radiofizika* **5**, 468.

(Paper 69D1-439)